Volume 10 Number 8 1982 Nucleic Acids Research

Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells

Melanie Ehrlich*, Miguel A.Gama-Sosa*, Lan-Hsiang Huang*, Rose Marie Midgett*, Kenneth C. Kuo+, Roy A.McCune+ and Charles Gehrke+

*Department of Biochemistry, Tulane Medical School, New Orleans, LA 70112, and +Department of Biochemistry, Experiment Chemical Laboratories, University of Missouri, Columbia, MO 65201, USA

Received 13 January 1982; Revised and Accepted 29 March 1982

ABSTRACT

Analysis of the total base composition of DNA from seven different normal human tissues and eight different types of homogeneous human cell populations revealed considerable tissue-specific and cell-specific differences in the extent of methylation of cytosine residues. The two most highly methylated DNAs were from thymus and brain with 1.00 and 0.98 mole percent 5-methylcytosine (m^5C), respectively. The two least methylated DNAs from in vivo sources were placental DNA and sperm DNA, which had 0.76 and 0.84 mole percent m^5C, respectively. The differences between these two groups of samples were significant with $p<0.01$. The m^5C content of DNA from six human cell lines or strains ranged from 0.57 to 0.85 mole percent. The major and minor base composition of DNA fractionated by reassociation kinetics was also determined. The distribution of m^5C among these fractions showed little or no variation with tissue or cell type with the possible exception of sperm DNA. In each case, nonrepetitive DNA sequences were hypomethylated compared to unfractionated DNA.

INTRODUCTION

5-Methylcytosine (m^5C) is present as a minor base in the DNA of all examined vertebrates and higher plants and is the only modified base naturally found in these DNAs (1). Recent studies indicate that for many genes associated with vertebrate development, there are tissue- or cell-specific differences in the extent of DNA methylation. Generally, but not always, an inverse correlation has been observed between methylation of certain cytosine (C) residues and transcriptional activity of DNA sequences (2-6). Restriction endonucleases were used in these studies to probe methylation of C residues in specific oligonucleotide sequences of a limited number of genes. In addition, from paper or thin layer chromatographic analyses of acid hydrolysates of DNA, Vanyushin et al.(7-9) and Kappler (10) reported evidence for tissue-specific differences in the total m^5C content of DNA. These studies of DNA digested to the free bases have been variously interpreted by other investigators (2,3,11-14) as showing either significant, insignificant, or questionable differences in the m^5C content

0305-1048/82/1008-2709$ 2.00/0

of the DNA from different animal tissues.

We have recently developed a high performance liquid chromatography (HPLC) system (15) for separation of the four major deoxynucleosides from each other and from the two minor naturally occuring deoxynucleosides, 5-methyldeoxycytidine (m^5dCyd) and N^6-methyldeoxyadenosine (m^6dAdo). The latter is found in DNA of many procaryotes and some invertebrates (1). In preparation for HPLC, we quantitatively digested the DNA with enzymes to deoxynucleosides. Our methodology allows accurate determination of the levels of m^5dCyd and, by inference, m^5C in as little as 1-5 µg of DNA containing only 0.5-1 mole percent m^5C residues (15). By this HPLC analysis we have unequivocably demonstrated tissue-specific differences in the overall extent of methylation of unfractionated human DNA and of human DNA sequences fractionated according to their rate of reassociation. Furthermore, the nature of these variations suggests that differentiation may involve considerable increases as well as decreases in DNA methylation.

Materials and Methods

Purification of DNA: The following cultured human cells were used as sources of DNA: Molt 4F, RPMI 1788, Raji, and IM-1, all non-virus producing, lymphoblastoid cell lines; HeLa, a carcinoma-derived cell line; and two normal skin fibroblast cell strains harvested after 13-20 cell population doublings. Of the lymphoblastoid cell lines, only Molt 4F was derived from cells which were not oncogenically transformed; only this line consisted of T cells. DNA was isolated from nuclei preparations as described previously (16). Radiolabeled DNA was obtained from approximately 1-6 x 10^7 cultured cells labeled with 1 mCi of [6-^{3}H]uridine, [^{3}H-methyl]methionine, or 0.2 mCi of [^{14}C-methyl]methionine in RPMI 1640 medium (GIBCO) with 40% of the standard methionine concentration (for Molt 4F) or BME medium (GIBCO) containing 1 mM sodium pyruvate (for fibroblasts or HeLa cells). Media was supplemented with dialyzed fetal calf serum to 10% and sodium formate to a final concentration of 1 mg/ml. The yields of methionine-labeled DNA, which was purified as described below, were 2 x 10^4 - 4 x 10^5 cpm; 98% of the radioactively labeled material was rendered acid-soluble by digestion with DNase I and <2% was digested under exhaustive conditions with RNase I or proteinase K (EM Biochemicals). Furthermore, >85% of the radioactivity from formic acid digests of these samples cochromatographed with m^5C (16).

In vivo derived sources of DNA were fresh autopsy specimens from accident victims, lymphocytes from blood of healthy individuals, and pooled nor-

mal sperm. Tissues were minced, homogenized, defatted by extraction with ether and ethanol, and then digested for 16 h at 60°C with proteinase K (50 µg/ml) in 10 mM Tris-HCl, 50 mM EDTA, 0.2 M NaCl, pH 7.5, with 0.5% sodium lauroyl sarcosinate. Sperm were lysed by the method of Borenfreund et al. (17). DNA was purified from these samples by extraction with freshly distilled phenol and chloroform/isoamyl alcohol (24/1, v/v), precipitation with ethanol by spooling (18), and then incubation of the DNA dissolved in 0.015 M NaCl, 0.0015 M sodium citrate, pH 7.3, with 50 µg/ml of RNase I and 12.5 units/ml of RNase T1 (Worthington Biochem. Corp.) for 14 h at 37°C. After further incubation with 50 µg/ml of proteinase K, the DNA was extracted with chloroform/isoamyl alcohol and precipitated as described above. The A_{260}/A_{280} ratio of all samples was 1.9.

Reassociation Kinetics: To reduce DNA to fragments of an average size of ∿0.3 kb, DNA solutions were sonicated in 0.05 M sodium phosphate buffer, pH 6.8, (PB) with 1 M NaCl after bubbling with nitrogen. The size range of the fragments was determined by polyacrylamide gel electrophoresis with HaeIII fragments of ØX174 RF as markers. Sonicates were chromatographed on Chelex-100 (Bio Rad), dialyzed against 0.01 M PB, heat denatured, quickly cooled to 60°C, and adjusted to 0.12 M PB. The DNA samples were then incubated to achieve a Cot of 0.05 molar sec and chromatographed on a hydroxyapatite column by standard techniques (19). If the Cot $\leq$0.05 fraction was to be further fractionated, it was denatured and reannealed in 0.05 M PB to a Cot of $<10^{-3}$ and chromatographed as above. The Cot >0.05 fraction was reassociated at a Cot of 50 and chromatographed on hydroxyapatite. In all cases the Cot value refers to the original concentration of DNA before fractionation and was corrected for losses during work-up (approximately 10%). When the concentration of the sodium phosphate buffer used for reannealing was changed from the standard 0.12 M, the Cot value was corrected as described by Britten et al. (19). DNA fractions were prepared for HPLC analysis by dialysis against 10 mM Tris-HCl, 1 mM EDTA, 0.3 M sodium acetate, pH 7.8, and then precipitation with ethanol.

Determination of nucleoside content of DNA: Unlabeled DNA samples (50 µl) in 10 mM-Tris-HCl, 0.1 mM EDTA, 4 mM $MgCl_2$, pH 7.6, were incubated with 40 µg/ml DNase I (EP; Sigma) for 16 h at 37°C. Then, sodium acetate, pH 5.3, $ZnSO_4$, and nuclease P1 (Boehringer-Mannheim) were added to final concentrations of 60 mM, 1 mM, and 30 µg/ml, respectively. After 8 h at 37°C, Tris-HCl, pH 8.4, and heat-treated (15) Escherichia coli alkaline

phosphatase (BAPC; Sigma) were added to final concentrations of 50 mM and 10 units/ml, respectively. Incubation at 37°C was continued for 16 h. These digests were directly analyzed for their deoxynucleoside composition by reversed phase, high performance liquid chromatography (HPLC) as described elsewhere (15 and Gehrke, McCune, Ehrlich, and Kuo, in preparation).

RESULTS

Tissue-specific or cell-specific differences in the m^5C content of human DNA

DNA samples from various normal tissues from 9 individuals were analyzed for their m^5C content by HPLC after quantitative hydrolysis to deoxynucleosides (15). Tissues were obtained from seven males and two females ranging in age from 8 months to 82 years. Despite their various ages and sexes, individuals did not differ in the extent of DNA methylation of a given tissue type or cell population at a significance level of $p<0.01$. For example, thymus DNA derived from two females, 8 and 12 months of age, was not significantly different from that of a 40 year old male. Heart DNA from three individuals was exceptional in displaying such differences with a $p<0.05$, but since only a total of seven determinations of three donors (males, 26-40 years of age) was involved and the differences in m^5C content were rather small (Table 1), these differences were probably insignificant. On the other hand, the m^5C content of DNA from different tissues or cell types varied considerably even when tissues from the same individual were compared. From 0.76 to 1.00 mole percent of the DNA bases were m^5C (Table 1). The greatest differences were seen in the m^5C content of placental or sperm DNA compared to that of brain or thymus DNA (Table 1). Even among adult organs, significant differences in the extent of methylation of C residues were observed (Table 1). The significance of many of these differences was established by Sheffe's method for comparison of means (20). For example, the m^5C content of the placental DNA samples was significantly different ($p<0.01$) from that of all the other examined cell populations except those of sperm and heart. A similar statistical analysis, revealed that the extent of methylation of sperm DNA was significantly different ($p<0.01$) from those of spleen, lymphocytes, thymus, and brain and that, at the same level of confidence, liver and brain DNAs had significantly different m^5C contents. Since all tissue homogenates and cell lysates had been treated extensively with proteinase K and sodium lauroyl sarcosinate in the presence of EDTA and 0.2 M NaCl before extraction with phenol (21) there should have been no significant, preferential losses of DNA sequences

Table 1. Mean m^5C levels in the DNA from various human tissues or cell populations.[a]

Source of DNA	Mean Mole %	Standard deviation	Number of individuals	Number of determinations
Placenta	0.76	0.03	6	22
Sperm	0.84	0.01	6[b]	17
Heart	0.87	0.03	3	7
Liver	0.88	0.02	9	18
Lungs	0.91	0.04	5	13
Spleen	0.93	0.03	7	16
Lymphocytes	0.96	0.01	2	3
Brain	0.98	0.03	5	13
Thymus	1.00	0.02	3	9

[a]The mean was based on the average of replicate determinations for each individual. Analysis was by HPLC at the deoxynucleoside level. Samples from each individual were analyzed at least in duplicate with the exception of one lymphocyte DNA sample.
[b]Six pooled sperm samples from approximately eight donors per sample were used and counted as individuals in this analysis.

during the extraction. Furthermore, we purified DNA from a nuclear fraction of placenta by chromatography on a hydroxyapatite column in the presence of 8 M urea without the use of organic solvents for extraction (22). The m^5C content of this sample (0.77 mole %) was essentially the same as that of parallel samples extracted by the standard procedure. We conclude that the differences in m^5C content among the various types of DNA samples was not an artifact of DNA isolation.

In the DNA digests of all of these tissues the only minor deoxynucleoside which was detected was m^5dCyd. Less than 0.02 mole percent of the nucleosides in these digests was m^6dAdo. There was no significant variation in the major base composition of the DNA samples. For example, the C content of the DNA from all examined tissues or in vivo-derived cell populations was constant within experimental error, namely, 20.89 mole % with a standard deviation (SD) of 0.51 mole %.

Extent of methylation of the DNA from cultured human cells

The m^5C content of different established human cell lines and of human diploid fibroblast cell strains was examined (Table 2). The lymphoblastoid cell lines IM-1 and RPMI 1788, have low levels of m^5C in their DNA compared

to those of most of the cell populations obtained without in vitro propagation, including lymphocytes (Table 1). There was considerable variation in the extent of DNA methylation in the four different lymphoblastoid cell lines (Table 2). The diploid skin fibroblast cell strain had a relative low level of m^5C compared to that of the in vivo-derived cells (Tables 1 and 2).

In one experiment the effects of cell division upon DNA methylation levels was examined. Two strains of skin fibroblasts were maintained at confluency for 35 days in Dulbecco's minimal essential medium with 1% fetal calf serum, 50 µg/ml gentamycin, and 5 mM HEPES buffer pH 7.2. Concurrently, identical cultures were induced to divide almost continuously by culturing the cells in the above medium containing 10% fetal calf serum and trypsinizing and subdividing the cell populations every week into 4-6 fresh bottles as the cultures approached confluency. One of the cell strains had a genomic m^5C content of 0.71 mole percent whether it was maintained at confluency, passed uninterruptedly, or maintained at confluency and then passed. The corresponding m^5C compositions for the other cell strain, 0.76, 0.72, and 0.72, respectively, were not significantly different. These DNA samples as well as those from the

Table 2. Content of m^5C in the DNA of various types of cultured human cells

Cell type	Mean mole % $m^5C \pm SD$[a]	$m^5C_{Cot} \div m^5C_{total}$[b]		
		Cot ≤0.05	Cot 0.05-50	Cot >50
Skin fibroblast	0.72 ± 0.01	1.68	1.18	0.72
HeLa	0.70 ± 0.10	1.82	0.88	0.70
Molt 4F	0.84 ± 0.02	2.23	1.16	0.66
RPMI 1788	0.57 ± 0.006	3.05	0.86	0.72
Raji	0.85 ± 0.005	-	-	-
IM-1	0.68 ± 0.009	-	-	-

[a]The mean mole % m^5C was determined as described in Table 1. Three to seven determinations were made on at least two different batches of cells.

[b]The enrichment of m^5C in the RPMI 1788 DNA fractions was determined from unlabeled DNA as the mole % of m^5dCyd in the given fraction ÷ mole % m^5dCyd in the total DNA from the cell population from which the fraction was derived. The analogous ratio for skin fibroblast, HeLa and Molt 4F DNA was determined from radioactive DNA samples as the % of [^{14}C-met]DNA sequences in the fraction ÷ the % of [^{3}H-Pyr]DNA sequences in the fraction.

examined cell lines had essentially the same C content as did DNA from human tissues.

The intragenomic distribution of m^5C among human DNA sequences of different degrees of repetitiveness

DNA fragments of ∿0.3 kb from various types of human tissue or cell samples were fractionated by reassociation kinetics (19,23) and then analyzed for their extent of methylation. The first set of experiments utilized DNA from cultured cell lines or strains, which had been labeled selectively in the m^5C residues with [^{14}C-methyl]methionine or [^{3}H-methyl]methionine or equally in all three pyrimidine residues with [6-^{3}H]uridine. A comparison of the percentage of radioactivity in the Cot ≤0.05 fraction from [^{3}H-Pyr]DNA and that from the homologous [^{3}H-m^5C]DNA or [^{14}C-m^5C]DNA revealed that these sequences had ∿70-120 % more m^5C than did total DNA whereas m^5C residues were underrepresented in the single copy fraction (Cot >50) fraction (Table 2).

By HPLC, we analyzed the total base composition of unlabeled fractionated DNA from RPMI 1788 cells and from nine different types of *in vivo*-derived tissues or cell populations. The average mole percentages of C ± SD in the Cot ≤0.05, 0.05-50, and >50 fractions from all the *in vivo*-derived DNA samples were 20.6 ± 0.6, 20.0 ± 0.05, and 19.3 ± 0.7, respectively; the average guanine (G) contents were, 21.8 ± 1.7, 21.2 ± 1.1, 20.4 ± 2.1 mole %. Analogously, slightly more thymine and adenine were found in the higher Cot fractions than in the lower Cot fractions. In contrast to the very small differences in the major base composition, the m^5C content of the fractions varied much. Cot ≤0.05 fractions from various *in vivo* sources contained ∿1.4-1.85 times as much m^5C as did unfractionated DNA whereas the Cot >50 fractions had ∿0.7-0.8 times as much (Table 3).

The variation from tissue to tissue in the extent of enrichment or deficiency in m^5C in a given Cot fraction (mole % m^5dCyd in fraction ÷ mole % m^5dCyd in the total DNA) may be mostly the result of experimental error due, for example, to differences in the percentage of the DNA sequences found in a given Cot fraction (Table 3). For example, in four Cot ≤0.05 samples of liver DNA, the enrichment in m^5C ranged from 1.55 to 1.83. Similarly, in one Cot >50 sample from brain DNA, there was 0.74 times as much m^5C as in unfractionated DNA, whereas another analogous sample had 0.85 times the m^5C content of total DNA. However, from triplicate Cot ≤0.05 and Cot >50 fractions from sperm DNA, the relative standard deviation was only 2%. Therefore, based on the data in Table 3, sperm DNA seems to be

Table 3. Distribution of m^5C among human DNA sequences fractionated by reassociation kinetics.[a]

Source of DNA	Cot ≤0.05 fraction			Cot 0.05-50 fraction			Cot >50 fraction		
	Mole % m^5C	% of genome[b]	$\frac{m^5C_{Cot}}{m^5C_{total}}$[c]	Mole % m^5C	% of genome	$\frac{m^5C_{Cot}}{m^5C_{total}}$	Mole % m^5C	% of genome	$\frac{m^5C_{Cot}}{m^5C_{total}}$
Placenta	1.26	24	1.63	0.77	13	1.00	0.56	63	0.74
Sperm	1.20	15	1.43	0.90	25	1.07	0.71	60	0.84
Heart	1.49	19	1.71	0.83	25	0.95	0.58	56	0.67
Liver	1.47	19	1.67	0.99	15	1.12	0.62	66	0.71
Lung	1.61	20	1.85	1.04	22	1.20	0.69	58	0.79
Spleen	1.62	24	1.80	-	-	-	-	-	-
Lymphocytes	1.59	22	1.66	1.13	10	1.17	0.67	68	0.71
Brain	1.52	15	1.58	1.06	18	1.09	0.76	67	0.79
Thymus	1.58	15	1.62	1.17	11	1.20	0.78	74	0.80

[a]Unlabeled DNA digests were analyzed by HPLC at the deoxynucleoside level. The values shown for liver and brain are the mean of four and two determinations, respectively, on two different DNA samples. The results for sperm DNA are the averages from three different DNA samples. All other data are from only a single sample.

[b]The percentage of the genome in a given fraction was determined by absorption of 260 nm light and corrected for hypochromicity. The observed differences from sample to sample in the percentage of DNA sequences reannealing at a given Cot are due to experimental variation and are not significant.

[c]The mole % of m^5dCyd in the given fraction ÷ mole % of m^5dCyd in the total DNA from the tissue or cell population from which the fraction was derived.

preferentially hypomethylated in its repetitive DNA sequences.

Five DNA samples from various somatic tissues were further fractionated into Cot $<10^{-3}$, 10^{-3}-0.05, 0.05-1, 1-50, 50-500 and >500 fractions with an average of 4, 14, 7, 4, 4, and 67 % of the DNA sequences, respectively. The mean mole percentages of m^5C in these DNA fractions were as follows: 1.34, 1.51, 1.02, 0.99, 0.67, and 0.66, respectively, with relative standard deviations of 13-22%. The average major base composition of these different subfractions varied much less than the m^5C content did. For example, for C the mole percentages ± SD of these subfractions were: 20.1 ± 0.8, 21.0 ± 0.7, 20.3 ± 0.4, 20.3 ± 0.8, 19.2 ± 0.7, and 19.5 ± 0.9, respectively.

One placental DNA sample was treated with S1 nuclease to remove single-stranded tails from partially double-stranded DNA after hydroxyapatite chromatography. The m^5C content of the Cot $<10^{-3}$ (foldback) and 10^{-3}-0.05 (highly repetitive) fractions were 1.30 and 1.39 mole percent, respectively, with S1 treatment followed by ethanol precipitation and 1.24 and 1.28 mole percent, without S1 treatment. Radiolabeled fibroblast and Molt 4F DNA samples were reannealed at a Cot of 10^{-4}. The resulting Cot$<10^{-4}$ fractions also had less hypermethylation than did the Cot 10^{-4}-0.05 fractions.

DISCUSSION

Previous studies of the DNA of various animals indicated that there were small tissue-specific differences in the m^5C content (7-10). Those experiments involved hydrolysis of DNA to the free bases with concentrated acid, which can cause significant and variable amounts of m^5C residues in DNA to be deaminated to T residues (24). Quantitation of the bases separated by paper or thin layer chromatography was from the $A_{290} - A_{310}$ of eluted spots (7-9) or from their radioactivity due to *in vivo* incorporation of [2-^{14}C]deoxycytidine (10). Problems of base trailing, lack of sensitivity, and requirement for *in vivo* radiolabeling have been eliminated by a highly sensitive and accurate HPLC analysis of enzymatic digests of DNA (15, Gehrke *et al.*, in preparation). With this methodology we have proven that significant differences can be found in the percentage of C residues which are methylated in various human tissues (Table 1). Furthermore, individual differences in the m^5C content of a given tissue were found to be negligible. The DNA used for these analyses was total cellular DNA. Since mitochondrial DNA constitutes less than 1% of the genome (25) and the observed tissue-specific differences im m^5C content were found in each

examined reassociation subfraction of the DNA (Table 3), these differences cannot be attributed to variation in the amount of organelle DNA.

Tissue or cell specificity in DNA methylation patterns has been demonstrated for a number of structural genes (2-6, 26-28) as well as for ribosomal RNA genes (29). In these studies hypomethylation is generally associated with gene expression. This tissue specificity of DNA methylation is not confined to a relatively small number of genes because our results indicate that one cell from a given human tissue can have millions or tens of millions more of its DNA cytosine residues methylated than does another type of cell from the same individual. Therefore, theories of how changes in the extent of DNA methylation arise during development will have to address massive differences in the number of methylated cytosine residues per genome as well as tissue-specific, transcription-associated variations in methylation of specific genes. The relative rate of cell turnover is probably not generally a major determinant of the extent of DNA methylation since in a number of the examined tissues there is no correlation between these cell renewal rates (30,31) and the levels of m^5C in the DNA (Table 1). For example, heart with its low rate of cell turnover has a low m^5C content compared to that of various other adult somatic tissues; whereas, brain, which also has only a small percentage of dividing cells, has a high m^5C content (Table 1).

Also, available data does not indicate a simple correlation between the extent of transcription of the genome and the overall levels of DNA methylation. Studies of the complexity of nuclear or cellular RNA from mice, rats, and rabbits indicate that in brain a much higher percentage of single copy sequences is transcribed than in liver (32-36). A similar situation seems to apply to RNA synthesis in human brain and liver (32). As discussed above, transcription of specific genes has been associated with hypomethylation; nonetheless, human brain DNA is 11% more methylated than human liver DNA (Table 1). This suggests that the observed tissue- and cell-specific differences in the genomic m^5C content are not due to generalized changes in DNA methylation programmed for control of transcription of a portion of the genome and limited to that fraction of the genome. Rather, this tissue- and cell-specific variation in the total m^5C content of DNA may be independent of hypomethylation of certain highly transcribed genes.

A limitation in the analysis of tissues is that they are complex populations of cells. Given the heterogeneity of the cells in the tissues

examined in this study, it is remarkable that we observed as large differences in overall levels of m^5C deoxynucleotide residues as we did (Table 1). Fractionated cell populations derived from the above tissues might reveal yet more heterogeneity in the amount of DNA methylation. The relatively homogenous cell populations which we examined, sperm, lymphocytes, and six types of cell cultures (Table 1 and 2), displayed significant differences in the m^5C content of their DNA. Differences in the percentage of C residues which are methylated have also been found in various murine cell lines (10,37). These differences could reflect either changes occurring during in vitro propagation (including preferential loss of part of the genome) or heterogeneity in the m^5C content of the original in vivo populations of cells.

By analysis of human DNA sequences fractionated according to their rate of reassociation, we have shown that highly repetitive DNA sequences (Cot 10^{-3}-0.04) and foldback DNA sequences (Cot $<10^{-3}$) have a considerably higher portion of their C residues methylated than do single copy or moderately repetitive DNA sequences, regardless of the cells of origin (Tables 2 and 3). Similar findings have been made in other types of mammalian tissues (9) and cultured cells (12,38). However, unlike the results reported by Romanov and Vanyushin for bovine DNA (9), the tissue-specific differences in overall levels of m^5C in human DNA were approximately evenly spread throughout the genome, except in the case of sperm DNA (Table 3), and were not limited to the repetitive DNA sequences. Sturm and Taylor (39) found that bovine sperm DNA contains 50% less m^5C than does calf thymus DNA and from studies of a bovine satellite DNA, they postulated that hypomethylation of satellite DNA could be the main factor accounting for the differences in overall DNA methylation. Our data indicate that hypomethylation of human sperm DNA was due to decreases in the m^5C content of single copy sequences and middle repetitive sequences as well as in the highly repetitive fraction. Although we observed less of a decrement in the m^5C content of the single copy fraction of sperm DNA than in the other fractions, it is noteworthy that any decrease in these sequences was observed since a number of genes in sperm DNA have been shown to be more methylated or at least as methylated as those same genes in various somatic tissues (3,4,26,40). Our results and those on methylation of bovine sperm DNA (39) and the activation of the heterchromatic X chromosomes in female mammalian cells (41,42) suggest that demethylation of certain DNA sequences might occur during gametogenesis.

That human sperm DNA has a lower level of methylation than does DNA

from a number of somatic tissues (Tables 1 and 2) implies that at some stages in human development there are increases in the extent of DNA methylation in addition to the previously observed decreases in methylation of specific genes (3,4,26). From comparisons of methylation of certain highly transcribed genes in various vertebrate cells (3,4,26,40), it has been proposed that hypomethylation may be necessary for turning on expression of certain genes (1) involved in development. However, as documented in this study for the genome at large, and in previous studies (3,4,43,44) for specific vertebrate genes, tissue-specific differences in DNA methylation, which cannot be correlated with the transcriptional activity are also found. It is not yet known, whether the tissue-specific variation which we observe in the overall extent of methylation of highly repetitive, moderately repetitive, and single copy human DNA sequences is a byproduct or determinant of differentiation.

ACKNOWLEDGEMENTS

We are very grateful to Drs. Monroe Samuels, Joseph Lanasa, Larry Wilson and Mrs. Lilly Dunn for providing us with various human cell samples. Dr. Frances Mather kindly helped us with the statistical analysis. This study was supported in part by National Institutes of Health Grant CA-19942 to M.E. and C. W. G.

REFERENCES

1. Ehrlich, M. and Wang, R. Y.-H. (1981) Science 212, 1350-1357.
2. McGhee, J. D., Ginder, G. D. (1979) Nature 280, 419-420.
3. Mandel, J. L. and Chambon, P. (1979) Nucleic Acid Res. 7, 2081-2103.
4. van der Ploeg, L.H.T. and Flavell, R. A. (1980) Cell 19, 947-958.
5. Shen, C. K. J., and T. Maniatis, (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 6634-6643.
6. Compere, S. J. and Palmiter, R. D. (1981) Cell 25, 233-240.
7. Vanyushin, B. F., Tkacheva, S. G., and Belozersky, A. N. (1970) Nature 225, 948-949.
8. Vanyushin, B. F., Mazin, A. L., Vasilyev. V. K., and Belozersky, A. N. (1973) Biochim. Biophys. Acta 299, 397-403.
9. Romanov, G. A. and Vanyushin, B. F. (1981) Biochim. Biophys. Acta 653, 204-218.
10. Kappler, J. W. (1971) J. Cell. Physiol. 78, 33-36.
11. Rubery, E. D., and Newton, A. A. (1973) Biochim. Biophys. Acta 324, 24-36.
12. Sawecka, J., Kornacka, L. and Malec, J. (1978) Experientia 35, 1166-1167.
13. Lapeyre, J. N. and Becker, F. F. (1979) Biochem. Biophys. Res. Commun. 87, 698-705.
14. Razin, A. and Riggs, A. D. (1980) Science 210, 604-610.
15. Kuo, K. C., McCune, R. A., Gehrke, C. W., Midgett, R., Ehrlich, M. (1980) Nucleic Acids Research 8, 4763-4776.

16. Ehrlich, M., Sarafyan, L. P., Simpson, N., and Downing, A. (1978) Biochim. Biophys. Acta 517, 43-54.
17. Borenfreund, E., Fitt, E., and Bendich, A. (1961) Nature 191, 1375-1377.
18. Marmur, J. (1961) J. Mol. Biol. 3, 206-218.
19. Britten, R. J., Graham, D. F., Neufeld, B. R. (1974) in Methods in Enzymology, L. Grossman and K. Moldave, Eds., Vol. 29E, pp. 363-405. Academic Press, New York.
20. Sheffe, H. (1959) The Analysis of Variance, Wiley, New York.
21. Rae, P. M. M., Barnett, T. R., and Babbitt, D. G. (1976) Biochim. Biophys. Acta 432, 154-160.
22. Meinke, W., Goldstein, D. A., and Hall, M. R. (1974) Anal. Biochem. 58, 82-88.
23. Schmid, C. W. and Deininger, P. L. (1975) Cell 5, 345-358.
24. Ford, J. P., Coca-Prados, M., and Hsu, M.-T. (1980) J. Biol. Chem. 255, 7544-7547.
25. Bogenhagen, D. and Clayton, D. A. (1974) J. Biol. Chem. 249, 7991-7995.
26. Jones, R. E., DeFeo, D., and Piatigorsky, J. (1981) J. Biol. Chem. 256, 8172-8176.
27. Groudine, M., Eisenman, R., and Weintraub, H. (1981) Nature 292, 311-317.
28. Yagi, M. and Koshland, M. E. (1981) Proc. Natl. Acad. Sci. USA 78, 4907-4911.
29. Bird, A. P., Taggert, M. H., and Gehring, C. A. (1981) J. Mol. Biol. 152, 1-17.
30. Cameron, I. L. (1970) Texas Rep. Biol. Med. 28, 201-248.
31. Edwards, J. L. and Klein, R. E. (1961) Am. J. Path. 38, 437-451.
32. Markov, G. G., Yaneva, J. N., Markova, N. G., and Ivanov, I. G. (1981) Int. J. Biochem. 13, 121-124.
33. Brown, I. R. and Church, R. B. (1972) Devel. Biol. 29, 73-84.
34. Hahn, W. E. and Laird, C. D. (1971) Science 173, 158-161.
35. Grouse, L., Chilton, M.-D., and McCarthy, B. J. (1972) Biochemistry 11, 799-805.
36. Chikaraishi, D. M., Deeb, S. S., and Sueoka, N. (1978) Cell 13, 111-120.
37. Fabricant, J. D., Wagner, E. F., Auer, B., and Schweiger, M. (1979) Exp. Cell Res. 124, 25-29.
38. Drahovsky, T., Boehm, T. L. J., and Kreis, W. (1979) Biochim. Biophys Acta 563, 28-35.
39. Sturm, K. S. and Taylor, J. H. (1981) Nucleic Acids Res. 9, 4536-4546.
40. Bird, A., Taggert, M., and MacLeod, D. (1981) Cell 26, 381-390.
41. Mohandas, T., Sparkes, R. S., and Shapiro, L. J. (1981) Science 211, 393-396.
42. Kratzer, P. G. and Chapman, V. M. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 3093-3097.
43. Waalwijk, C. and Flavell, R. A. (1978) Nucleic Acids Res. 5, 4631-4640.
44. Bird, A. P. and Southern, E. M. (1978) J. Mol. Biol. 118, 27-47.